

GWDG  
**NACHRICHTEN**  
06-07|13

Projektmanagement

.....  
New scientific  
compute cluster

.....  
Erfolgreiche  
ISO-9001-Zertifizierung

.....  
Verteilte Datenspeicherung  
mit Ceph

ZEITSCHRIFT FÜR DIE KUNDEN DER GWDG

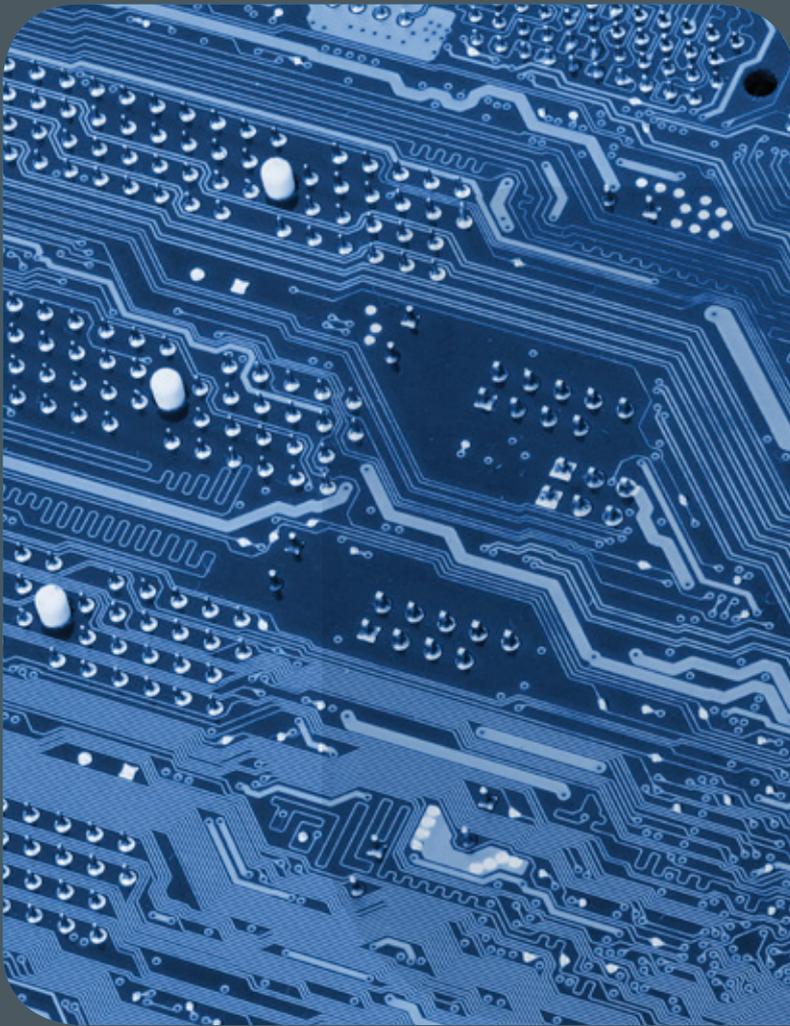


**PROJEKTE**

**PROJEKTE**

**PROJEKTE**

 **GWDG**  
Gesellschaft für wissenschaftliche  
Datenverarbeitung mbH Göttingen



## GWDG NACHRICHTEN

# 6-7|13 Inhalt

.....

**4 Neuer Dienst bei der GWDG: Werkzeuge für das Projektmanagement** **6 New scientific compute cluster available at GWDG**  
**8 GWDG erfolgreich nach ISO 9001:2008 zertifiziert** **9 Verteilte Datenspeicherung mit Ceph** **12 Kurz & knapp** **13 Stellenangebot**  
**14 Kurse**

### Impressum

.....  
Zeitschrift für die Kunden der GWDG

ISSN 0940-4686  
36. Jahrgang  
Ausgabe 6-7/2013

**Erscheinungsweise:**  
monatlich

[www.gwdg.de/gwdg-nr](http://www.gwdg.de/gwdg-nr)

**Auflage:**  
500

**Fotos:**  
© dp@pic – Fotolia.com (1)  
© DOC RABE Media – Fotolia.com (4)  
© contrastwerkstatt – Fotolia.com (13)  
© MPLbpc-Medienservice (3)  
GWDG (2, 6, 14)

#### Herausgeber:

Gesellschaft für wissenschaftliche  
Datenverarbeitung mbH Göttingen  
Am Faßberg 11  
37077 Göttingen  
Tel.: 0551 201-1510  
Fax: 0551 201-2150

**Redaktion:**  
Dr. Thomas Otto  
E-Mail: [thomas.otto@gwdg.de](mailto:thomas.otto@gwdg.de)

**Herstellung:**  
Maria Geraci  
E-Mail: [maria.geraci@gwdg.de](mailto:maria.geraci@gwdg.de)

**Druck:**  
GWDG / AG H  
E-Mail: [printservice@gwdg.de](mailto:printservice@gwdg.de)



Prof. Dr. Ramin Yahyapour  
ramin.yahyapour@gwdg.de  
0551 201-1545

## *Liebe Kunden und Freunde der GWDG,*

*es ist geschafft! Wie in den GWDG Nachrichten 04/2013 berichtet, stand nach langer Vorbereitung die erste Auditierung nach ISO 9001 an. Der Erfolg belohnt die nicht unerheblichen Anstrengungen aller Beteiligten und liefert damit die Grundlage für einen kontinuierlichen Qualitätsmanagementprozess. Es liegt in der Natur der Sache, dass dieser nie abgeschlossen sein wird und ständig weiterentwickelt werden muss. Eine Ausweitung auf IT-Sicherheit nach ISO 27001 ist sinnvoll und daher von uns mittelfristig angestrebt.*

*Im Juni startete auch der allgemeine Produktionsbetrieb für den ersten Teil unseres neuen Compute Clusters. Gleichzeitig wurden die Details für die Konfiguration der zweiten Phase definiert, die noch in diesem Jahr realisiert werden soll. Hierbei möchte ich Sie auch auf das neue Kursangebot zur Parallelverarbeitung hinweisen. Die Heranführung an das wissenschaftliche Computing ist eine wichtige Aufgabe, um aktuelle und künftige Generationen von Rechnersystemen effizient zu nutzen. Da immer mehr wissenschaftliche Disziplinen rechen- und datenintensive Anwendungen erfordern, sind entsprechende Kompetenzen zunehmend wichtig. Dem wollen wir auch mit erweiterten Kursangeboten Rechnung tragen.*

*In dieser Ausgabe finden Sie nähere Details zu diesen und weiteren Themen. Ich wünsche Ihnen viel Freude beim Lesen.*

**Ramin Yahyapour**

GWDG – IT in der Wissenschaft



## Neuer Dienst bei der GWDG: Werkzeuge für das Projektmanagement

Text und Kontakt:  
Maik Srba  
maik.srba@gwdg.de  
0551 39-21108

Die GWDG führt zu einer Reihe eigener Projekte durch und ist zudem an einer Reihe nationaler und internationaler Forschungsprojekte beteiligt. Dabei spielt das Projektmanagement eine wichtige Rolle. Für die Verwaltung solcher Projekte wurde ein Projektmanagementsystem installiert. Die Erfahrungen, die dabei gesammelt wurden, sollen auch den Nutzern der GWDG zugute kommen. Zu diesem Zweck wurde ein Projektmanagementdienst ins Leben gerufen, der allen Nutzern der GWDG zur Verfügung steht. Der Dienst ist in die Benutzerverwaltung der GWDG integriert.

### KERNFUNKTIONEN

Mit **Chiliproject** bietet die GWDG einen Dienst für das Verwalten von Projekten an. Dieser Dienst bietet ein professionelles Tool für die Verfolgung Ihrer Projekte und deren Anforderungen. Die Projektteilnehmer sind damit in der Lage, ihre Aufgaben zu koordinieren und die Ergebnisse nachvollziehbar zu dokumentieren. Der Projektmanagementdienst erlaubt es, den Ablauf des Projektes zu analysieren und dadurch frühzeitig Probleme zu erkennen. Dies ermöglicht eine gezielte Steuerung des Projektes, um Risiken zu verringern. Des Weiteren ermöglicht er, die Aufgaben eines Projektes kollaborativ zu bewältigen, Dokumentationen zu erstellen, Ergebnisse zu teilen und den Source Code zu verwalten. Dazu bietet der Projektmanagementdienst eine Reihe von Features an, die je nach Bedarf aktiviert oder deaktiviert werden können (siehe Tabelle).

### New GWDG service: Tools for project management

Also in science project management has become an integral part. Besides working on urgent scientific issues, new discoveries, papers, posters, or Ph.D. theses are still other things to do and it should not be neglected. Important conferences must be kept in mind, work has to be organized, submissions must be prepared and valuable data are managed. Most often this is handled through emails, Excel lists, list notes or USB sticks. These informations are scattered and mostly not available when needed. Also such information is not easy to share and complicates working. The project management service at GWDG helps you to manage your important information and data centrally, share results easier, work together on tasks and manage your source code.

## Projektverwaltung

- Organisieren von Projekten und Unterprojekten
- Gantt-Diagramm zeigt Fortschritt und Aufgabenverlauf
- Kalender zeigt wichtige Termine und Deadlines
- Erzeugen von Issues für wichtige Termine und Aufgaben
- E-Mail-Benachrichtigungen
- Berichte für Fortschritt und Verantwortlichkeiten
- Organisieren der Aufgaben in Kategorien und Releases
- Übersicht über die Projektaktivitäten
- Erzeugen einer Roadmap und Release-Planung
- Selbstdefinierte Filter für die Aufgabenübersicht

## Benutzerverwaltung

- GWDG-Account-Integration
- Definieren von Benutzerrollen
- Verwaltung von eigenen SSH-Keys für den Zugriff auf das Quellcode-Management

## Kollaboratives Arbeiten

- Projektinternes Forum für Diskussionen
- Nachrichten, um Teammitglieder zu informieren
- Wiki, um Informationen zu dokumentieren und zu teilen
- Teilen von Dokumenten und Dateien
- Kommentieren von Aufgaben
- RSS-Feeds, um informiert zu bleiben

## Quellcode-Management

- Git und SVN zum Verwalten des Quellcodes
- Integration mit der Aufgabenliste
- Selbstdefinierbare post-receipts für Git
- Automatische Spiegelung für Git
- Zugang über benutzereigene SSH-Keys
- Deployment SSH-Keys
- WebDav-Zugang
- Webbrowser-Ansicht

## Sicherheit & Backup

- Backup für Projekte- und Quellcode-Verwaltung
- Zugang über HTTPS
- GWDG-gehostet

Der Projektmanagementdienst ist durch weitere, frei zur Verfügung stehende Plugins erweiterbar und er steht als Open

Source zur Verfügung. Der Projektmanagementdienst der GWDG ist damit flexibel und kann an veränderte Bedingungen des Projektmanagements angepasst werden. Um langfristig den Erhalt der eigenen Projektdaten zu gewährleisten, halten wir uns Migrationen der Daten zu anderen Systemen offen. Der Projektmanagementdienst beschränkt sich nicht auf die Bereitstellung von Chili-project als Dienst, sondern geplante Zusatzfunktionen gehen über eine einfache Projektverwaltung hinaus. So wollen wir nicht nur bei der Verwaltung der Projekte helfen, sondern planen eine Integration in unsere GWDG Compute Cloud. Durch die Integration in die GWDG Compute Cloud sind automatische Tests und ein Aufbau eines eigenen Continuous-Integration(CI)-Systems wie Jenkins CI möglich.

Der Dienst ist in die Benutzerverwaltung der GWDG integriert und kann von allen bei der GWDG registrierten Personen genutzt werden. Neben der kontinuierlichen Verbesserung des Dienstes durch weitere nützliche Tools, die den Projektablauf unterstützen, ist die Integration in das Portal der GWDG vorgesehen. Damit wird das Ziel verfolgt, den Projektmanagementdienst für alle bei der GWDG registrierten Benutzer als Self-Service zur Verfügung zu stellen. Der Dienst kann individuell angepasst werden, so dass verschiedene Projektstrukturen unterstützt werden. Neben klassischen Projekten werden auch agile Projekte unterstützt. Der Dienst kann in Form eines einfachen Git oder SVN-Repositories oder einer einfachen Projektverwaltung bis hin zu einem umfangreichen Projektmanagement mit integriertem Testworkflow verwendet werden.

## SOURCE CODE MANAGEMENT

Mit dem integrierten Source Code Management (SCM) bietet der Dienst die Möglichkeit, SVN und Git innerhalb eines Projektes einsetzen, um die Implementierung der für das Projekt benötigten Software zu verwalten. Sobald man für den Projektmanagementdienst freigeschaltet wurde, stehen diese beiden SCMs als Self-Service dem Nutzer zur Verfügung. Durch das Abschalten von Features im Projektmanagementsystem kann die Source-Code-Verwaltung auch separat genutzt werden. Jedes Git-Repository kann über eine selbst definierte Post-Receipt-URL mit einem Web Server kommunizieren. Sobald ein Push auf das Repository erfolgt wird die Post-Receipt Url informiert. Solche Post-Receipt-URLs können beispielsweise benutzt werden, um Continuous Integration Builds anzustoßen oder Code auf einen produktiven Server auszurollen. Der Zugriff auf das Git-Repository kann optional über http(s) oder SSH geschehen. Durch Hinterlegen eines öffentlichen Schlüssels im Projektmanagementdienst kann der Zugriff passwortfrei geschehen. ●

### Infobox

Der Projektmanagementdienst der GWDG ist im WWW unter <https://projects.gwdg.de> zu finden. Er kann per E-Mail an [projects-admins@gwdg.de](mailto:projects-admins@gwdg.de) beantragt werden.



The new cluster hardware, situated at the Laserlabor Göttingen facility

# New scientific compute cluster available at GWDG

## Text und Kontakt:

Dr. Christian Boehme  
christian.boehme@gwdg.de  
0551 201-1839

Tim Ehlers  
tim.ehlers@gwdg.de  
0551 201-1520

The first installment of our new scientific compute cluster, announced in the February issue of “GWGD-Nachrichten”, is now available to all compute cluster users at GWDG. Originally planned for mid April, the start of the production phase had been delayed by supply shortages and unforeseen integration problems. However, after successful being stress tested by the users of the beta phase for four weeks, the system is now ready for general use. The 5,440 cores of the first installment already surpass the total computing power available before its introduction by a wide margin, and therefore will alleviate the long waiting times currently observed on our clusters.

## ACCESS

For using our compute cluster you need an GWDG account. This account is, by default, not activated for the use of the compute resources. To get it activated, please send a formless email to [support@gwdg.de](mailto:support@gwdg.de).

Once you gain access, you can login to the fronted nodes [gwdu101.gwdg.de](http://gwdu101.gwdg.de) and [gwdu102.gwdg.de](http://gwdu102.gwdg.de). These nodes are accessible via ssh from within the GÖNET. If you come from the internet, first login to [login.gwdg.de](http://login.gwdg.de). From there you can then reach the frontends.

The frontends are meant for editing, compiling, and interacting with the batch system, but please don't use them for testing for more than a few minutes, since all users share resources on the frontends and will be impaired in their daily work, if you overuse them. [gwdu101](http://gwdu101) is an AMD based system, while [gwdu102](http://gwdu102) is Intel based. If your software takes advantage of special CPU dependent features, it is recommended to use the same CPU architecture for compiling as targeted for running your jobs.

## PREPARING BINARIES

Most of the third-party software installed on the cluster is not located in the default path. To use it, the corresponding *module* must be loaded. Furthermore, through the module system you can setup environment settings for your compiler to use special

libraries. You can find a list of installed modules, sorted by categories, by entering *module avail* on one of the frontends [gwdu101](http://gwdu101) or [gwdu102](http://gwdu102). The command *module list* gives you a list of currently loaded modules.

The big advantage of this system is the (relative) simplicity with which one can coordinate environment settings, such as *PATH*, *MANPATH*, *LD\_LIBRARY\_PATH* and other relevant variables, dependent on the requirements of the use-case.

## Neuer Scientific Compute Cluster bei der GWDG verfügbar

Der erste Teil unseres neuen, in der Februar-Ausgabe der GWGD-Nachrichten angekündigten Scientific Compute Cluster ist jetzt für alle Compute-Cluster-Nutzer verfügbar. Der ursprünglich für Mitte April geplante Start des Produktionsbetriebs hatte sich auf Grund von Lieferschwierigkeiten und unvorhergesehenen Problemen bei der Integration verzögert. Nach einem vierwöchigen Lasttest durch die Nutzer der Beta-Phase ist der Rechner aber nun bereit für den allgemeinen Nutzerbetrieb. Bereits die 5.440 Rechenkerne der aktuellen Lieferphase übertreffen die Gesamtleistung der vorher vorhandenen Systeme deutlich. Die zuletzt beobachteten langen Wartezeiten auf Rechenzeit sollten sich daher erheblich verkürzen.

To use a module, you can explicitly load the version you want with `module load software/version`. If you leave out the version, the default version will be used. Logging off and back in, will unload all modules, as will `module purge`. You can unload single modules by entering `module unload software`.

The recommended compiler module for C, C++, and Fortran code is `intel/compiler/64/13.0`. We also provide Gnu and Open64 compilers, the PGI compiler suite will follow later this year. Open64 is often recommended for AMD CPUs, but we do not have experience with it. For math (*BLAS* and *fftw3*) the Intel MKL is a good default choice (`intel/mkl/64/11.0`), with ACML being an alternative for AMD processors. Usually it is not necessary to use *fftw3* modules alongside with the MKL, as the latter provides *fftw* support as well. Please note that the module `python/scipy/mkl/0.12.0` provides Python's *numpy* and *scipy* libraries compiled with Intel MKL math integration, thus offering good math function performance in a scripting language.

`intel-mpi/64/4.0.3` and the various OpenMPI flavors are recommended for MPI, mostly due to the fact that the *mvapich* and *mvapich2* libraries lack testing. Please note that for HDF5 (and therefore also for parallel NetCDF) library usage it is required to use the Intel MPI Beta with module `intel-mpi/64/4.1.0/024-beta`, as previous Intel MPI versions fail several of the library checks.

## RUNNING JOBS

Job scheduling is handled by IBM LSF, as on our older resources. In most cases this means that only few changes to your current job scripts and `bsub` command lines are necessary. Presently, only the new cluster can be used from the new frontends, the older resources will be integrated later. There are three different node types in the cluster, each being served by one or more queues of the general form `<nodetype>[-<category>]`. For example, *fat* type nodes are served by the *fat* queue for normal jobs, and by *fat-long* for long jobs.

In the first installment there are 160 *mpi* nodes with 16 Intel CPU Cores and 64 GByte of memory, 20 *fat* nodes with 64 AMD CPU cores and 256 GByte of memory, and 25 *tp* (for *throughput*) nodes with 64 AMD cores and 64 GByte of memory. The *mpi* nodes provide the bulk of the compute power of our new cluster, and are meant for all types of applications. However, they are especially well suited for large MPI jobs, as they have a balanced compute to network performance ratio. The *fat* nodes are meant for shared memory parallelized workloads scaling beyond 16 cores, and for all applications requiring more than 64 GByte of memory on a single node. The *tp* nodes are not meant for general usage, as they are not owned by GWDG.

As stated before, `bsub` is used to submit jobs to the cluster. For example, to avoid running large tests on the frontend (a good idea!) you can get an interactive session (with the bash shell) on one of the *mpi* nodes with

```
bsub -ISs -q mpi-short -n 16 /bin/bash
```

You will get a shell prompt, as soon as a suitable node becomes available. Single thread, non-interactive jobs can be submitted with

```
bsub -q mpi ./myexecutable
```

Note that a single thread job submitted like this will share its execution host with other jobs. It is therefore expected that it does not use more than the memory available per core! On both

the *fat* and *mpi* nodes this amount is 4 GByte. If your job requires more, you must assign it additional cores. For example, if your single thread job requires 128 GByte of memory, you must submit it like this:

```
bsub -q fat -n 32 ./myexecutable
```

Shared memory parallelized jobs can be submitted with

```
bsub -q mpi -n 16 -R ,span[hosts=1] -a openmp ./myexecutable
```

The `span` option is required, as, without it, LSF will assign cores to the job from several nodes, if that is advantageous from the scheduling perspective. This behavior is different to the setup of our older resources.

OpenMPI jobs can be submitted as follows:

```
bsub -q mpi -n 256 -a openmpi mpirun.lsf ./myexecutable
```

For Intel MPI jobs it suffices to use `-a intelmpi` instead of `-a openmpi`. Please note that LSF will not load the correct modules (compiler, library, MPI) for you. You either have to do that before executing `bsub`, in which case your setup will be copied to the execution hosts, or you will have to use a job script and load the required modules there.

A new feature in LSF is *pinning* support. *Pinning* (in its most basic version) means instructing the operating system to not apply its standard scheduling algorithms to your workloads, but instead keep processes on the CPU core they have been started on. This may significantly improve performance for some jobs, especially on the *fat* nodes with their high CPU core count. *Pinning* is managed via the MPI library, and currently only OpenMPI is supported. There is not much experience with this feature, so we are interested in your feedback. Here is an example:

```
bsub -R „span[ptile=16] affinity[core(1):cpubind=core]” -q mpi -n 256 -a openmpi mpirun.lsf ./myexecutable
```

Besides the options shown in this article, you can of course use the options for controlling walltime limits (`-W`), output (`-o`), and your other requirements as before. You can also continue to use job scripts instead of the command line (with the known `#BSUB <option> <value>` syntax). Please consult the documentation on our webpages, or the LSF man pages if you need further information.

## STORAGE

The standard UNIX home directories are available on the new cluster. However, there is also a parallel file system available as `/scratch`. This file system is especially well suited for workloads. It is recommended that you create your own user directory under `/scratch`, and run all your jobs from there, in order to get the best possible performance. Backup is not available for `/scratch`, so, albeit the system is secured against data loss due to typical hardware failures, indispensable data should be copied to your home directory. We will provide some automated means to do this at a later point in time.

Beyond this common storage area there is also a small (80 GByte) temporary file space mounted locally on each node as `/local`. This file space is realized by fast SSD drives, so it is well suited for *scratch* data, if one does not need more than the available space.

This document is also hosted on our website, and new features and information will be added, as they become available. Further sources of information are the man pages of LSF, and of the libraries and compilers. As usual, you can reach us at [support@gwdg.de](mailto:support@gwdg.de). ■

# GWDG erfolgreich nach ISO 9001:2008 zertifiziert

## Text und Kontakt:

Dr. Thomas Otto  
thomas.otto@gwdg.de  
0551 201-1828

Die GWDG hat in diesen Tagen nach erfolgreichem Zertifizierungsaudit von der DQS GmbH, Deutsche Gesellschaft zur Zertifizierung von Managementsystemen, das Zertifikat nach ISO 9001:2008 erhalten. Damit ist die GWDG eines der ersten wissenschaftlichen Rechenzentren, dem nach einer umfassenden und unabhängigen Begutachtung das Vorhandensein eines Qualitätsmanagementsystems gemäß den Forderungen der ISO 9001:2008 bescheinigt wird. Weitere Schritte für Zertifizierungen nach zusätzlichen IT-spezifischen Normen sind mittel- und langfristig geplant.



Für die GWDG als langjähriger IT-Dienstleister mit einem breiten Leistungsspektrum für die Wissenschaft gehören die ständige Qualitätssicherung und kontinuierliche Qualitätsverbesserung zu ihren wichtigsten Zielen, um die Anforderungen der Kunden nach einer leistungsfähigen und innovativen IT-Infrastruktur auch erfolgreich erfüllen zu können. Mit der jetzt zertifizierten erfolgreichen Einführung und Anwendung eines Qualitätsmanagementsystems wird ein bedeutender Schritt zur Erreichung dieser Ziele gemacht.

Dokumentierte Prozesse und ein zugehöriges Monitoring sind wichtige Voraussetzungen für ein erfolgreiches Qualitätsmanagement. Nachdem in der GWDG das zugehörige Projekt fundiert und umfangreich vorbereitet worden war, begutachtete ein Auditor der DQS im Laufe des Aprils eingehend die Prozesse der GWDG, zu denen die Kernprozesse Support und IT-Beratung, Administration und Systempflege, Backup und Archiv, Institutsbetreuung und Kundens Schulung sowie weitere Führungs- und Unterstützungsprozesse gehören. „Die Einführung eines Qualitätsmanagementsystems

war für uns eine große Herausforderung, der wir uns gerne gestellt haben, und ein gutes Stück Arbeit, die sich aber gelohnt hat“, sagt Prof. Dr. Ramin Yahyapour, Geschäftsführer der GWDG. „Unsere eigene Leistungsfähigkeit sowie die Ausrichtung unseres Managementsystems auf die ständige Verbesserung der Servicequalität für die Kunden hin werden jetzt auch durch eine externe Zertifizierung dokumentiert“, so Yahyapour weiter.

Damit die Positionierung im wissenschaftlichen Umfeld gefestigt sowie ein hochwertiger und professioneller IT-Betrieb, wie ihn die Kunden der GWDG schon seit vielen Jahren gewohnt sind, auch in Zukunft gesichert werden, sind weitere Zertifizierungen geplant. Mittelfristig ist die Integration der Informationssicherheit nach ISO 27001 in das Managementsystem angedacht, und langfristig soll dieses dann um die Aspekte des IT-Servicemanagements nach ISO 20000 ergänzt werden. „Mit dem ISO-9001-Zertifikat haben wir den Grundstein für den schrittweisen Ausbau unseres Managementsystems und zukünftige weitere IT-spezifische Zertifizierungen gelegt. Die Anforderungen unserer zahlreichen Kunden hinsichtlich IT-Sicherheit und IT-Service werden weiter zunehmen, so dass den entsprechenden Zertifizierungen hierzu große Bedeutung zukommen wird“, so Yahyapour abschließend. ■

## GWDG successful certified by ISO 9001:2008

GWDG has now received the certificate ISO 9001:2008 after a successful certification audit by DQS GmbH. Thus the GWDG is one of the first scientific computing centres, to which was certified the presence of a quality management system in accordance with the requirements of ISO 9001:2008 after a comprehensive and independent review. Additional steps for further certifications to IT-specific standards are planned medium and long term.

# Verteilte Datenspeicherung mit Ceph

## Text und Kontakt:

Piotr Kasprzak  
piotr.kasprzak@gwdg.de  
0551 201-1579

Exponentielles Datenwachstum stellt traditionelle Client-Server-Architekturen zur Speicherung von Daten zunehmend in Frage. Das verteilte Dateisystem „Ceph“, welches seit einigen Jahren als Open-Source-Projekt für Linux entwickelt wird, hat sich um Ziel gesetzt, Skalierungsgrenzen bei Massenspeichersystemen aufzulösen und eine Plattform bereitzustellen, die Tausenden von Klienten den Zugriff auf ExaByte an Daten ermöglichen kann und als Hardwaregrundlage Standard-PCs nutzt. Doch Ceph bietet neben dem Fokus auf Skalierbarkeit noch mehr: Alternativ zum dateorientierten Zugriff ermöglicht Ceph auch, auf den verwalteten Massenspeicher block- und objektbasiert zuzugreifen. Hochverfügbarkeit und intelligente Mechanismen zum Datenmanagement runden das Gesamtpaket ab.

## EINLEITUNG

Die heutige Zeit ist gekennzeichnet durch eine Explosion der zu speichernden und verwaltenden Datenmengen. Anwendergenerierte Inhalte für „Social Networks“, mit dem Handy aufgenommene Filme, aber auch von technischen Geräten und Sensoren generierte Daten (z. B. Gensequenzen) werden dafür sorgen, dass weiterhin mit Wachstumsraten von ca. 40 % pro Jahr gerechnet werden muss (siehe <http://www.emc.com/collateral/analyst-reports/idc-the-digital-universe-in-2020.pdf>). Der bisherige Ansatz des Scale-Up, also des Ausbaus der bestehenden zentralen Systeme durch die Hinzunahme weiterer Ressourcen (wie z. B. mehr Speicher oder schnellere CPUs) funktioniert aufgrund finanzieller, aber zunehmend auch technischer Schranken oft nicht mehr. Ähnlich dem Paradigmenwechsel bei CPUs, wo zusätzliche Performance nicht mehr durch die Steigerung der Leistung der einzelnen Verarbeitungseinheit erreicht wird, sondern durch eine größere Anzahl von Verarbeitungseinheiten und Parallelität, muss in Zukunft auch beim Aufbau und der Verwaltung von Massenspeicher das Prinzip des Scale-Out angewendet werden, also die Verteilung der Last bei Zugriff und der Gesamtkapazität des Massenspeichers auf eine Menge von Knoten. Genau hier setzt Ceph an.

## CEPH

Ceph entstand als Resultat der Promotion von Sage Weil (<http://ceph.com/papers/weil-thesis.pdf>), dem Initiator des Projekts, Kernentwickler und Gründer von Inktank (<http://www.inktank.com>), einem Start-Up, das dem üblichen Geschäftsmodell von Open Source Software folgend Dienstleistungen und professionellen Support rund um Ceph anbietet.

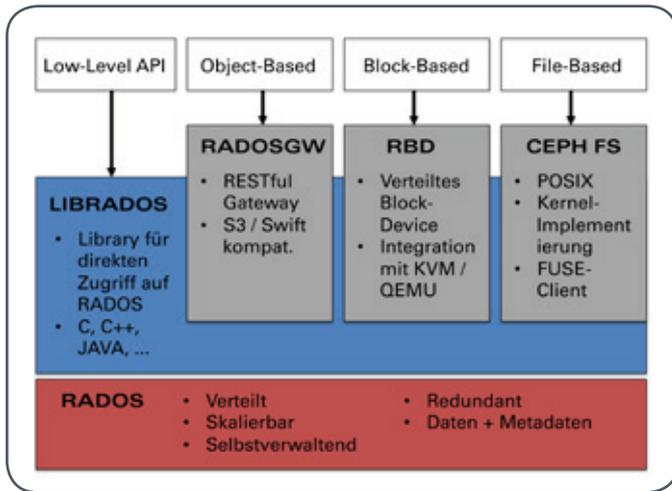
Als Hardwaregrundlage nutzt Ceph Standard-PCs, an die einfach die gewünschte Anzahl von Festplatten angeschlossen wird. RAID-Controller sind nicht notwendig: Ceph sorgt für die

Datensicherheit selbst, indem eine von dem Benutzer spezifizierbare Anzahl von Replika vorgehalten wird, die bei Ausfall einzelner Knoten oder Festplatten zur Rekonstruktion der jeweiligen Daten herangezogen werden. Das System kann problemlos im laufenden Betrieb durch Hinzunahme weiterer Knoten vergrößert oder bei Bedarf auch verkleinert werden und balanciert sich selbstständig entsprechend vorgegebener Regeln aus.

Die logische Architektur von Ceph ist in Abbildung 1 dargestellt. Im Zentrum steht der RADOS Object Store („Reliable, Autonomic Distributed Object Store“). Dieser stellt das eigentliche Massenspeicher-Backend und entsprechende Abstraktionen und Primitiven zur Verfügung, auf denen die weiteren Komponenten von Ceph wie das POSIX-konforme verteilte Dateisystem CephFS fußen. Die primitiven Ceph-Objekte und assoziierte Metadaten werden dabei einfach als Dateien in den mit einem lokalen Dateisystem formatierten Datenbereichen des jeweiligen Knotens abgelegt, die für die Verwendung durch Ceph vorgesehen sind. Als lokale Dateisysteme werden dabei *btrfs*, *ext4* und *xfs* direkt unterstützt. Obwohl *btrfs* einige für den Einsatz unter Ceph

## Distributed storage system Ceph

Exponentiell data growth challenges traditionell client server architectures for storage systems. The distributed storage system „Ceph“ which is being developed as Open Source software for Linux targets the current scalability limits of storage systems and offers a platform which aspires to be able to handle the access of thousands of clients to exabyte of data while running on cheap commodity hardware. Additional features include support for block and object storage from the same storage platform (also called „unified storage“), high availability, redundancy and intelligent mechanisms for self management.



1\_Architektur von Ceph

interessante Eigenschaften besitzt (Copy-on-Write, Checksumming, Snapshots) und Ceph über entsprechende Integrationspunkte verfügt, um diese nutzen zu können, wird aufgrund der immer noch vorhandenen Instabilität von *btrfs* zu *xf*s für den produktiven Einsatz geraten (<http://ceph.com/docs/master/rados/configuration/filesystem-recommendations/>). In Zukunft könnte auch „ZFS on Linux“ (<http://zfsonlinux.org/>) eine gute Alternative sein – entsprechende Anpassungsarbeiten laufen gerade.

Eine Kerninnovation von Ceph und Mitgrund für die exzellente Skalierbarkeit liegt in der Art und Weise, wie Klienten auf Speicherobjekte zugreifen. Traditionelle verteilte Systeme nutzen hierfür häufig zentrale Server, die von den Klienten zunächst kontaktiert werden müssen, um den konkreten Speicherort der Daten zu bestimmen oder über die gar auch der gesamte Datenverkehr zu den eigentlichen Massenspeicherknoten abgewickelt werden muss. Bei Ceph dagegen findet diese Kommunikation direkt zwischen Klient und Massenspeicherknoten statt, ohne den Einsatz einer Zwischenschicht. Über den sogenannten CRUSH-Algorithmus ist jeder Ceph-Klient in der Lage, aus der Struktur des Clusters und einer Parametrisierung der Abbildung (CRUSHMAP), den Ablageort eines Datums und seiner Replika zu berechnen. Die Parametrisierung dieser pseudorandomisierten Abbildung kann genutzt werden, um spezifische, an die lokale physikalische Infrastruktur angepasste Datenplatzierungsregeln zu definieren. So kann beispielsweise angegeben werden, dass die Replika von bestimmten Daten jeweils nur in Knoten abgelegt werden, die sich in unterschiedlichen Racks befinden, das wichtige Daten häufiger repliziert oder performancerelevante Daten vorzugsweise auf SSDs abgelegt werden sollen.

```
rule data {
    ruleset 0
    type replicated
    min_size 2
    max_size 2
    step take rack1
    step chooseleaf firstn 0 type host
    step emit
}
```

2\_Eine generische CRUSH-Regel für die zweifache Replikation von Daten

Abbildung 2 zeigt als Ausschnitt aus einer CRUSHMAP eine generische Regel für die zweifache Replikation von Daten, durch

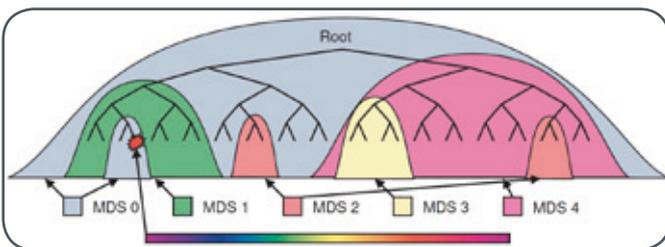
die zur Selektion von Zielknoten für die Datenablage als Einstiegspunkt in einen Teilbaum zur hierarchischen Abbildung der Infrastruktur, der hier nicht dargestellt ist, das Objekt *rack1* gewählt wird, welches einem bestimmten Rack mit Massenspeicherknoten entsprechen kann. Anschließend wird eine der definierten Anzahl von Replika (= 2) entsprechende Menge von Knoten bestimmt, die zur Ablage des Datums und seiner Replika genutzt werden soll. Die einzelnen Regeln definieren über die verwendeten Ids (*ruleset 0*) ein Regelwerk, welches den logischen Massenspeicheraggregations- und Verwaltungsabstraktionen in Ceph, den „Pools“, zugeordnet werden kann. Letztendlich ist es die Aufgabe der CRUSHMAP, zum einen möglichst genau die dem Massenspeicheraufbau zugrundeliegende physikalische Organisationsstruktur wie Standorte, Räume, Racks und Knoten hierarchisch abzubilden, und zum anderen Regeln zu definieren, wie Daten und ihre Replika innerhalb dieser Hierarchie sinnvoll abgelegt werden sollen.

Während der RADOS Object Store die notwendigen Basisfunktionalitäten für konsistenten Zugriff, Replikation und Ausfallerkennung und Behandlung implementiert, werden die für den Endanwender nutzbaren Zugriffsschnittstellen auf den Massenspeicher von separaten Komponenten zur Verfügung gestellt, die für den Zugriff auf RADOS eine entsprechende Bibliothek, die LIBRADOS, nutzen, die interessanterweise auch unabhängig eingesetzt werden kann, um unter Nutzung der bereitgestellten Primitiven auch andere Speicherkonzepte zu realisieren, wie beispielsweise solche, die auf spezifische Applikationen zugeschnitten sind. Bindings für verschiedene Programmiersprachen sind vorhanden.

Über das sogenannte RADOS-Gateway (RADOSGW) stellt Ceph „Object-Storage“ bereit, objektbasierten Speicher, auf den über HTTP zugegriffen werden kann, und bietet hierfür zwei unterschiedliche APIs für Endanwender an, die Kompatibilität zum einen mit Amazons S3, zum anderen mit OpenStack Swift ermöglichen. Der von Ceph mit RADOS bereits intern implementierte Object-Storage ist wesentlich primitiver und nicht für die direkte Nutzung durch Endanwender vorgesehen. Das RADOS-Gateway überbrückt das Funktionsdelta, indem es unter anderem den externen HTTP-Zugriff und Authentifizierungs- und Autorisierungsfunktionalitäten bereitstellt.

Ceph ermöglicht mit dem RADOS Block-Device (RBD) Zugriff auf verteilten blockbasierten Massenspeicher, der insbesondere hinsichtlich der Verwendung im Kontext von Virtualisierung ausgerichtet ist. So existiert eine direkte Integration mit QEMU / libvirt, so dass virtuelle Maschinen, die z. B. KVM als Hypervisor nutzen, auf einem RBD Block-Device laufen können. Hinzu kommt auch Unterstützung für RBD in Open Source „Infrastructure-as-a-Service“-Cloud-Lösungen wie CloudStack und OpenStack. Letztere ermöglicht die Verwendung von RBD auch zur Ablage von Abbildern von Festplatten und Snapshots von virtuellen Maschinen (Treiber für OpenStack „Glance“) oder die Einbindung von zusätzlichem Block-Storage in virtuelle Maschinen (Treiber für OpenStack „Cinder“). Das RBD kann natürlich auch direkt in Linux als Block-Device genutzt werden. Die notwendige Linux-Kernel-Unterstützung ist seit v2.6.37 vorhanden und RBD seitdem Bestandteil des Standard-Kernels. RBD unterstützt auch verschiedene moderne Features zur Verwaltung von blockbasiertem Massenspeicher wie „Thin Provisioning“ (Speicher wird erst allokiert, wenn er tatsächlich genutzt wird), Snapshots und Vergrößerung oder Verkleinerung.

CephFS ist das eigentliche verteilte Dateisystem von Ceph. Es ist seit v2.6.34 Bestandteil des Standard-Linux-Kernels. Neben POSIX-Konformität nutzt es die Möglichkeiten des RADOS Object-Stores, um Skalierbarkeit bis hin zu Tausenden von Knoten, Hochverfügbarkeit und Ausfallsicherheit, n-fache Replikation sowie einfaches Management durch intelligente Selbstreorganisation bei z. B. Änderungen der Größe des zur Verfügung stehenden Massenspeichers zu ermöglichen. Dateien werden jeweils in Blöcke fester Größe unterteilt, die dann auf unterschiedlichen Massenspeicherknoten (sogenannten OSDs – „Object Storage Devices“) abgelegt werden. Dieses sogenannte „Striping“ ermöglicht nicht nur eine hohen Durchsatz und gute Skalierbarkeit, da beim Zugriff Ressourcen einer Anzahl von Knoten parallel genutzt werden können, sondern ermöglicht in Verbindung mit Replikation auch eine schnelle Regeneration bei Ausfall einzelner Knoten. Eine weitere interessante Eigenschaft von CephFS ist die Fähigkeit, je nach Last, nicht nur einzelne Verzeichnisse verschiedenen Metadaten-Knoten (MDS – „Metadata Server“) zwecks Lastausgleichs zuzuweisen, sondern auch große Verzeichnisse als solche entsprechend fragmentieren zu können. Dies ist in Abbildung 3 dargestellt. Der Klienten-Zugriff auf CephFS ist nicht nur über den Linux-Kernel möglich, sondern es existiert auch eine Userspace-Implementierung für FUSE.



3\_Ceph verteilt Teilbäume des Dateisystems dynamisch auf einzelne Metadaten-Server (MDS). Bei hoher Last können auch einzelne Verzeichnisse fragmentiert werden. (Quelle: Weil, Sage A. et. al.: Ceph: A scalable, high-performance distributed file system. In: Proceedings of the 7th symposium on operating systems design and implementation (OSDI'06), 2006, Seattle, Washington, pp. 307-320)

Während die Object-Storage-Komponente von Ceph, das RADOS-Gateway, sowie das verteilte Block-Devices, RBD, als stabil und produktionsreif angesehen werden, was sich unter anderem auch darin äußert, dass die hinter Ceph stehende Firma Inktank hier entsprechende Support-Dienstleistungen anbietet, ist dies für CephFS leider noch nicht der Fall. Zwar wird dieses durchaus schon produktiv eingesetzt und bei unseren internen Tests sind bisher keine Probleme aufgetreten, jedoch scheint es für eine generelle Empfehlung für einen Produktiveinsatz noch zu früh zu sein.

## INSTALLATION EINES CEPH-CLUSTERS

Das Aufsetzen eines Ceph-Clusters für Testzwecke erweist sich als relativ einfach (<http://ceph.com/docs/master/start/quick-start/>). Es existieren Skripte und Tools, die nach einigen einfachen Vorarbeiten (wie der Bereitstellung eines Zugangs per SSH zu den einzelnen Knoten), einen vollständigen Ceph-Cluster automatisch installieren können. In der aktuellen Stable-Release von Ceph, Cuttlefish (v0.61), wurde das bisher für diesen Zweck genutzte Tool `mkcephfs` (<http://ceph.com/docs/master/rados/deployment/mkcephfs/>) von `ceph-deploy` (<http://ceph.com/docs/master/rados/deployment/preflight-checklist/>) abgelöst, welches leider noch die

```
[global]
auth supported = cephx
auth cluster required = cephx
auth service required = cephx
auth client required = cephx

public network = 192.168.0.0/24
cluster network = 10.2.0.0/24

[osd]
osd journal size = 1000
# uncomment the following line if you are mounting with
# ext4
# filestore xattr use omap = true

# uncomment if you want journal on tmpfs
# journal dio = false

[mon.a]

host = mon-a
mon addr = 192.168.0.1:6789

[mds.a]
host = mds-a

[osd.0]
host = osd-0
public addr = 192.168.0.8
cluster addr = 10.2.0.8

[osd.1]
host = osd-1
public addr = 192.168.0.9
cluster addr = 10.2.0.9

[osd.2]
host = osd-2
public addr = 192.168.0.10
cluster addr = 10.2.0.10

[osd.3]
host = osd-3
public addr = 192.168.0.11
cluster addr = 10.2.0.11
```

### 4\_Konfigurationsdatei „ceph.conf“ für einen einfachen Ceph-Cluster

eine oder andere Macke hat, wenn man von den Standardfällen abweicht. Natürlich kann `mkcephfs` auch weiterhin für Ceph Cuttlefish genutzt werden. Abbildung 4 zeigt die zentrale Konfigurationsdatei `ceph.conf` für eine Beispielformatierung eines Ceph-Cluster, die unterschiedliche Netzwerke für die Kommunikation mit Klienten und Replikationsverkehr zwischen den einzelnen OSDs definiert und als Ausgangspunkt für eine manuelle Konfiguration des Clusters genutzt werden kann. Der einfache Aufbau von `mkcephfs` als Bash-Skript erlaubt es, dieses in diesem Fall auch als Dokumentation der notwendigen manuellen Schritte zu verwenden.

## FAZIT

Ceph stellt einen interessanten Ansatz dar, die sich zunehmend verschärfenden Skalierungs-, Management- und Verfügbarkeitsproblematiken bei Massenspeicher zu adressieren. Als „Unified Storage“ kann Ceph alle gängigen Anwendungsfälle, unabhängig

davon, ob datei-, block- oder objektbasierter Zugriff benötigt wird, von einer zentralen Plattform aus bedienen und bietet in den jeweiligen Teilbereichen zusätzlich eine Reihe innovativer Features. Auch wenn es noch vielleicht etwas zu früh ist, alle Komponenten von Ceph für einen produktiven Einsatz zu empfehlen, ist es sicherlich sinnvoll, die weitere Entwicklung zu beobachten. ■

# Kurz & knapp

## Kontingenzzuweisung für das dritte Quartal 2013

Die nächste Zuweisung von Institutskontingenten für die Inanspruchnahme von Leistungen der GWDG erfolgt am Montag, dem 1. Juli 2013. Die Höhe der Kontingente wird den Instituten per Brief oder per E-Mail mitgeteilt. Die Bemessung der Institutskontingente erfolgte nach den Vorläufigen Richtlinien des Beirats der GWDG und den Ergänzungen der Beiratskommission für die Verteilung von IT-Leistung entsprechend dem Verbrauch im Zeitraum vom 01.12.2012 bis 31.05.2013. Nicht verbrauchte Kontingente werden zu 50 % in das nächste Quartal übertragen. Negative Verbrauchswerte werden zu 100 % mit dem neuen Institutskontingent verrechnet.

Jeder Benutzer kann den aktuellen Stand des Institutskontingents durch die Eingabe des Kommandos *kontingent* auf einer Workstation des UNIX-Clusters oder im WWW unter <http://www.gwdg.de/index.php?id=1678> abfragen. Dort besteht auch die Möglichkeit, Informationen über den Stand des separaten Druckkontingents abzurufen.

Falls in Ausnahmefällen das Institutskontingent nicht ausreichen sollte, können begründete Anträge über <http://www.gwdg.de/index.php?id=799> gestellt werden. Solche Anträge sollen bis zum **19.08.2013** eingereicht werden.

Glässer

## Erweitertes Kursprogramm für die Nutzung der Rechencluster der GWDG

Für die effiziente Nutzung der von der GWDG für die Wissenschaft bereitgestellten Rechencluster bzw. Parallelrechner, die in diesem Jahr umfangreich erweitert werden (siehe auch den entsprechenden Artikel in dieser Ausgabe), gibt es ein entsprechendes Kursprogramm mit Einführungs- und Vertiefungskursen, das jetzt erweitert wurde. Kurzfristig wurden der Kurs „Using the GWDG Scientific Compute Cluster – An Introduction“ (23.09.2013; 0,5 Tage) und der Kurs „Parallelrechnerprogrammierung mit MPI“ (23.09. – 25.09.2013; 2,5 Tage) ins laufende Kursprogramm aufgenommen. Nähere Informationen hierzu finden Sie unter <http://www.gwdg.de/kurse> in den entsprechenden Übersichten.

Neben diesen beiden Kursen, die zukünftig zweimal jährlich angeboten werden sollen, weil sie für die gesamte Nutzerschaft der GWDG-Rechencluster von Interesse sein können, gibt es noch drei weitere Kurse zu spezielleren Themen, deren Interessentenkreis nicht so genau voraussehbar ist, dass eine genaue Zeitplanung möglich wäre: „Parallelrechnerprogrammierung mit OpenMP“, „GPU-Programmierung mit CUDA“ und „Einführung in Unified Parallel C (UPC) und Co-array FORTRAN (CAF)“. Diese Kurse werden deshalb jeweils nach Bedarf abgehalten, wobei Zeit, Ort und z. T. auch inhaltliche Schwerpunkte in Abstimmung mit den Interessenten festgelegt werden. Die Bedarfsermittlung hierfür erfolgt zu einen durch Umfragen, zum anderen aber auch aufgrund der Nachfragen von interessierten Nutzern, die gerne an [support@gwdg.de](mailto:support@gwdg.de) gerichtet werden können und dann bei der Planung entsprechend berücksichtigt werden. Nähere Informationen zu den fünf Kursen sind unter <http://www.gwdg.de/index.php?id=2830> zu finden.

Otto

## Fünf neue RRZN-Handbücher verfügbar

Ab sofort sind bei der GWDG folgende fünf neuen RRZN-Handbücher verfügbar:

- Access 2013 – Grundlagen für Anwender
- Access 2013 – Grundlagen für Datenbankentwickler
- Excel 2013 – Fortgeschrittene Techniken
- Outlook 2013
- PowerPoint 2013 – Fortgeschrittene Techniken

Weitere Informationen zu den RRZN-Handbüchern, insbesondere zum Verkauf bei der GWDG, finden Sie unter <http://www.gwdg.de/index.php?id=615>.

Otto

## Doppelausgabe der GWDG-Nachrichten

In Zukunft sollen die GWDG-Nachrichten immer schon zu Beginn des jeweiligen Monats erscheinen. Aus diesem Grund gibt es diesmal eine Doppelausgabe 06-07/2013. Die Augustausgabe 08/2013 erscheint dann zu Beginn des August.

Otto

**Die GWGD sucht** ab dem 1. September 2013 zur Unterstützung der Arbeitsgruppe „Nutzerservice und Betriebsdienste“ eine

## Studentische Hilfskraft

mit bis zu 80 Stunden Beschäftigungszeit im Monat. Die Vergütung erfolgt entsprechend den Regelungen für Studentische/Wissenschaftliche Hilfskräfte.

### Aufgabenbereich:

- Mitarbeit bei der Installation und Betreuung von Windows-Arbeitsplatzrechnern in den Instituten
- Mitarbeit in den Bereichen Sharepoint 2013 und Active Directory der GWGD
- Mitarbeit bei der Systemüberwachung und Peripheriebetreuung abends und an Wochenenden

### Anforderungen:

- Gute Kenntnisse in Windows-Betriebssystemen
- Schnelle Lernfähigkeit
- Gute Kommunikations- und Teamfähigkeit

Die GWGD will den Anteil von Frauen in den Bereichen erhöhen, in denen sie unterrepräsentiert sind. Frauen werden deshalb ausdrücklich aufgefordert, sich zu bewerben. Die GWGD ist bemüht, mehr schwerbehinderte Menschen zu beschäftigen. Bewerbungen Schwerbehinderter sind ausdrücklich erwünscht.

Wir bitten interessierte Damen und Herren um eine Bewerbung bis zum **31.07.2013** über das Online-Formular unter <http://www.gwdg.de/stellenangebote>.

## Stellenangebot

Fragen zur ausgeschriebenen Stelle beantworten Ihnen:

**Herr Eric Helmvoigt**

Tel.: 0551 201-1845

E-Mail: [eric.helmvoigt@gwdg.de](mailto:eric.helmvoigt@gwdg.de) oder

**Herr Stefan Quentin**

Tel.: 0551 201-1816

E-Mail: [stefan.quentin@gwdg.de](mailto:stefan.quentin@gwdg.de)



**INFORMATIONEN:**  
support@gwdg.de  
0551 201-1523

Juli bis  
Dezember 2013

# Kurse

KURS	VORTRAGENDE/R	TERMIN	ANMELDEN BIS	AE
EINFÜHRUNG IN WINDOWS 7	Buck	13.08.2013 9:00 – 12:30 und 13:30 – 15:30 Uhr	06.08.2013	4
INSTALLATION UND ADMINIS- TRATION VON WINDOWS 7	Buck	20.08.2013 9:00 – 12:30 und 13:30 – 15:30 Uhr	13.08.2013	4
DIE SHAREPOINT-UMGEBUNG DER GWDC	Buck	21.08.2013 9:00 – 12:30 und 13:30 – 15:30 Uhr	14.08.2013	4
ADMINISTRATION VON PCS IM ACTIVE DIRECTORY DER GWDC	Buck	27.08.2013 9:00 – 12:30 und 13:30 – 15:30 Uhr	20.08.2013	4
GRUNDLAGEN DER BILDBEAR- BEITUNG MIT PHOTOSHOP	Töpfer	04.09. – 05.09.2013 9:30 – 16:00 Uhr	28.08.2013	8
INDESIGN – GRUNDLAGEN	Töpfer	10.09. – 11.09.2013 9:30 – 16:00 Uhr	03.09.2013	8
GRUNKURS UNIX/LINUX MIT ÜBUNGEN	Hattenbach	17.09. – 19.09.2013 9:15 – 12:00 und 13:30 – 16:00 Uhr	10.09.2013	12
USING THE GWDC SCIENTIFIC COMPUTE CLUSTER	Dr. Boehme, Ehlers	23.09.2013 9:00 – 12:00 Uhr	16.09.2013	2
PARALLELRECHNERPRO- GRAMMIERUNG MIT MPI	Dr. Haan	23.09.2013 14:00 – 17:00 Uhr 24.09. – 25.09.2013 9:15 – 17:00 Uhr	16.09.2013	
OUTLOOK – E-MAIL UND GROUPWARE	Helmvoigt	26.09.2013 9:15 – 12:00 und 13:00 – 16:00 Uhr	19.09.2013	4

KURS	VORTRAGENDE/R	TERMIN	ANMELDEN BIS	AE
PHOTOSHOP FÜR FORTGESCHRITTENE	Töpfer	01.10. – 02.10.2013 9:30 – 16:00 Uhr	24.09.2013	8
CLIENT-MANAGEMENT MIT BARAMUNDI	Becker, Körmer, Quentin, Rosenfeld	17.10.2013 9:00 – 12:30 und 13:30 – 15:30 Uhr	10.10.2013	4
INDESIGN – AUFBAUKURS	Töpfer	22.10. – 23.10.2013 9:30 – 16:00 Uhr	15.10.2013	8
UNIX FÜR FORTGESCHRITTENE	Dr. Sippel	04.11. – 06.11.2013 9:15 – 12:00 und 13:15 – 15:30 Uhr	28.10.2013	12
EINFÜHRUNG IN DIE STATISTISCHE DATENANALYSE MIT SPSS	Cordes	13.11. – 14.11.2013 9:00 – 12:00 und 13:00 – 15:30 Uhr	06.11.2013	8
EINFÜHRUNG IN DAS IP-ADRESSMANAGEMENTSYSTEM DER GWDG FÜR NETZWERKBEAUFTRAGTE	Dr. Beck	28.11.2013 10:00 – 12:00 Uhr	21.11.2013	2
UNIX/LINUX-ARBEITSPLATZRECHNER – INSTALLATION UND ADMINISTRATION	Gedes, Dr. Heuer, Körmer, Dr. Sippel	02.12. – 03.12.2013 9:15 – 12:00 und 13:30 – 16:00 Uhr	25.11.2013	8
UNIX/LINUX-SERVER – GRUNDLAGEN DER ADMINISTRATION	Gedes, Dr. Heuer, Körmer, Dr. Sippel	04.12. – 05.12.2013 9:15 – 12:00 und 13:30 – 16:00 Uhr	27.11.2013	8
UNIX/LINUX – SYSTEMSICHERHEIT FÜR ADMINISTRATOREN	Gedes, Dr. Heuer, Körmer, Dr. Sippel	06.12.2013 9:15 – 12:00 und 13:30 – 15:00 Uhr	29.11.2013	4
ANGEWANDTE STATISTIK MIT SPSS FÜR NUTZER MIT VORKENNTNISSEN	Cordes	11.12. – 12.12.2013 9:00 – 12:00 und 13:00 – 15:30 Uhr	04.12.2013	8

#### Teilnehmerkreis

Das Kursangebot der GWDG richtet sich an alle Mitarbeiterinnen und Mitarbeiter aus den Instituten der Universität Göttingen und der Max-Planck-Gesellschaft sowie aus einigen anderen wissenschaftlichen Einrichtungen.

#### Anmeldung

Anmeldungen können schriftlich per Brief oder per Fax unter der Nummer 0551 201-2150 an die GWDG, Postfach 2841, 37018 Göttingen oder per E-Mail an die Adresse [support@gwdg.de](mailto:support@gwdg.de) erfolgen. Für die schriftliche Anmeldung steht unter <http://www.gwdg.de/antragsformulare> ein Formular zur Verfügung. Telefonische Anmeldungen können leider nicht angenommen werden.

#### Kosten bzw. Gebühren

Unsere Kurse werden wie die meisten anderen Leistungen der GWDG in Arbeitseinheiten (AE) vom jeweiligen Institutskontingent abgerechnet. Für die Institute der Universität Göttingen und

der Max-Planck-Gesellschaft erfolgt keine Abrechnung in EUR.

#### Absage

Sie können bis zu acht Tagen vor Kursbeginn per E-Mail an [support@gwdg.de](mailto:support@gwdg.de) oder telefonisch unter 0551 201-1523 absagen. Bei späteren Absagen werden allerdings die für die Kurse berechneten AE vom jeweiligen Institutskontingent abgebucht.

#### Kursorte

Alle Kurse finden im Kursraum oder Vortragsraum der GWDG statt. Die Wegbeschreibung zur GWDG sowie der Lageplan sind unter <http://www.gwdg.de/lageplan> zu finden.

#### Kurstermine

Die genauen Kurstermine und -zeiten sowie aktuelle kurzfristige Informationen zu den Kursen, insbesondere zu freien Plätzen, sind unter <http://www.gwdg.de/kurse> zu finden.



Gesellschaft für wissenschaftliche  
Datenverarbeitung mbH Göttingen